PCT/GB 2004 / 0 0 2 5 8 9

GB0412589

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

RECD 13 JUL 2004
WIPO PCT

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed Andrew Gersey

Dated 1 July 2004

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
F
17 JUN 2003
NEWPORT

The Patent Office

1/77

17JUN03 E815553 1 D01886
P01/7700 0.00-0313995.3

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference

RD-ATC-33

2. Patent application number
(The Patent Office will fill in this part)

0313995.3

17 JUN 2003

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

Advanced Technologies (Cambridge) Limited
Globe House
1 Water Street
London WC2R 3LA

Patents ADP number *(if you know it)*

~~57476442002~~ 75 77356002

If the applicant is a corporate body, give the country/state of its incorporation

England & Wales

4. Title of the invention

Plant Limit Dextrinase Inhibitor Promoter

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Margot Ruth WALFORD
Patents Department
British American Tobacco
R&D Centre
Regents Park Road
Southampton SO15 8TL

Patents ADP number *(if you know it)*

07515778004

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

YES

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	19 ✓
Claim *(s)*	
Abstract	SV
Drawing *(s)*	21 + 21

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	4 sequence listing ✓

11. I/We request the grant of a patent on the basis of this application.

Signature M Reiner Diederich Date 16.6.03

12. Name and daytime telephone number of person to contact in the United Kingdom — Mr Reiner Diederich — 02380 793731

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

LIMIT DEXTRINASE INHIBITOR PROMOTER

1. FIELD OF INVENTION

The invention relates generally to a DNA sequence identified from *Hordeum vulgare* which acts as a tissue specific promoter. The invention further relates to the use of said DNA sequence to regulate exogenous gene expression in the endosperm and aleurone tissues of the seed of a plant.

2. BACKGROUND

The expression of a gene is dependent upon its DNA sequence being transcribed into RNA by the action of RNA polymerase. To achieve this, the RNA polymerase must recognise and attach to a region of DNA sequence located upstream of (i.e. 5' to) the gene coding sequence in order for transcription to be initiated. Such a region is termed the promoter of the gene. The intrinsic nature of the promoter sequence determines the circumstances and the manner in which the gene is expressed.

There are, broadly speaking, four types of promoters found in plant tissues; constitutive, tissue-specific, developmentally-regulated, and inducible/repressible.

A constitutive promoter directs the expression of a gene throughout the various parts of a plant continuously during plant development, although the gene may not be expressed at the same level in all cell types. Examples of known constitutive promoters include those associated with the cauliflower mosaic virus 35S transcript (Odell et al, 1985, Nature 313 810-812), the rice actin 1 gene (Zhang et al, 1991, Plant Cell 3 1155-1165) and the maize ubiquitin 1 gene (Cornejo et al, 1993, Plant Molec. Biol. 29 637-646).

A tissue-specific promoter is one which directs the expression of a gene in one (or a few) parts of a plant, usually throughout the lifetime of those plant parts. The category of tissue-specific promoter commonly also includes tissue-preferred or tissue-enhanced promoters which may also direct expression at some level in tissues other than the preferred tissue. Examples of tissue-specific promoters known in the art include those associated with the patatin gene expressed in potato tuber and the high molecular weight glutenin gene expressed in wheat, barley or maize endosperm.

A developmentally-regulated promoter directs a change in the expression of a gene in one or more parts of a plant at a specific time during plant development. The gene may be expressed in that plant part at other times during plant development at a different (usually lower) level, and may also be expressed in other plant parts.

An inducible promoter is capable of directing the expression of a gene in response to an inducer. In the absence of the inducer the gene will not be expressed. The inducer may act directly upon the promoter sequence, or may act by counteracting the effect of a repressor molecule. The inducer may be a chemical agent such as a metabolite, a protein, a growth regulator, or a toxic element, a physiological stress such as heat, wounding, or osmotic pressure, or an indirect consequence of the action of a pathogen or pest. A developmentally-regulated promoter might be described as a specific type of inducible promoter responding to an endogenous inducer produced by the plant or to an environmental stimulus at a particular point in the life cycle of the plant. Examples of known inducible promoters include those associated with wound response, such as described by Warner et al (1993), temperature response as disclosed by Benfey & Chua (1989), and chemically induced, as described by Gatz (1995).

A promoter sequence may comprise of a number of defined domains necessary for its function. The first of these defined domains comprises approximately 70 base pairs located immediately upstream of (5' to) the structural gene and forms the core promoter. The core promoter contains the CAAT and TATA boxes and defines the transcription initiation site for the structural gene. A series of regulatory sequences upstream of the core promoter constitute the remainder of the promoter sequence and determine the expression levels, the spatial and temporal patterns of expression, and the response to inducers. In addition some promoters contain sequence elements which act to enhance the level of expression, for example that from the pea plastocyanin promoter as described in International Patent Application, Publication No. WO 97/20056.

Genetic modification of plants depends upon the introduction of chimaeric genes into plant cells and their controlled expression under the direction of a promoter. Promoters may be obtained from different sources including plants, fungi, bacteria and viruses, and different promoters may work with different efficiencies in different tissues.

It may often be desirable to express introduced genes in a number of different tissues within a plant. For example, the expression of a resistance to a pathogen or pest, or tolerance to temperature extremes might be best expressed throughout all tissues in the plant. Similarly it might be desirable to ensure the expression of the transgenes at all times throughout the development of the plant. Also, a promoter which is expressed in a manner that is immune to the influence of inducers or repressors resulting from unforeseen environmental stimuli may also be useful to ensure the continued expression of a trait. For these purposes, the use of a "constitutive" promoter would be desirable. Examples of constitutive promoters include the

CaMv 35S promoter. For cereals the ubiquitin promoter is a constitutive promoter of choice (Christensen & Quail,1996)

However, in some instances it is more desirable to control the location of gene expression in a transgenic plant to enhance the effect of gene expression by ensuring that expression occurs preferentially in those tissues where the effect of the gene product is most efficacious. By the same argument, modulated expression can reduce potential yield loss by reducing the resource drain on the plant. Further advantages include limiting the expression of agronomically useful yet generally deleterious genes to specific tissues by localisation and compartmentalisation of gene expression in cases where the gene product must be restricted to, or excluded from, certain tissues. For example, anther specific expression of the *suc* inhibitor genes (Mariani et al, 1990, Nature **347** 737) has been used in male sterility systems, whereas expression in other parts of the plant would result in toxicity.

Examples of endosperm specific barley promoters include the high molecular weight glutenin promoter (Halford et al). However, there is a lack of suitable endosperm specific promoters.

3. SUMMARY OF THE INVENTION

The present invention is directed to a DNA sequence identified from *Hordeum vulgare* (barley).

The present invention provides a DNA sequence comprising a promoter sequence, the DNA sequence being selected from the group consisting of: (a) a DNA sequence, the sequence being known herein as SEQ. ID. No:1; (b) a DNA sequence comprising a portion of the sequence of the isolated DNA of (a), and being capable of regulating the expression of a gene; (c) a DNA sequence, which sequence is at least 95% , 90%, 85% 80%, 75% or 70% homologous to the DNA sequence of (a) or (b) above; (d) a DNA sequence, which sequence is at least 80% identical to the DNA sequence of (a) or (b) and which is capable of regulating the expression of a gene; (e) a DNA sequence comprising a portion of the sequences of the DNAs of any one of (a) - (d) above and being capable of regulating the expression of a gene; (f) a DNA sequence, which sequence is complementary to the DNA sequence of any one of (a) to (e) above.

Preferably the DNA sequence is capable of regulating expression of a gene which encodes a limit dextrinase inhibitor protein (LDI).

More preferably the DNA sequence regulates expression of the gene encoding the Limit dextrinase inhibitor-like protein in *Hordeum vulgare.* Most preferably the DNA sequence regulates expression of the gene encoding the Limit dextrinase inhibitor protein in the endosperm or aleurone tissues of developing seeds of *Hordeum vulgare.*

The DNA sequences referred to herein may be isolated DNA sequences or, alternatively, may be synthesised DNA sequences.

The present invention also provides a recombinant DNA comprising vector DNA and a DNA sequence as described in any one of (a) to (f) above. The recombinant DNA may suitably further comprise the DNA coding sequence of a gene.

Preferably the recombinant DNA resides in a host cell suitable for transcription and translation.

The present invention also provides a method of regulating the expression of a gene, the method utilising the DNA sequence as described in any one of (a) to (f) above operably associated with the coding sequence of a gene.

The present invention also provides a transgenic plant, the cells of which plant comprise the DNA as described in any one of (a) to (f) above operably associated with a gene coding sequence.

The present invention further provides a method of modifying the metabolism within the cells of a transgenic plant utilising the DNA sequence as described in any one of (a) to (f) above operably associated with the coding sequence of a gene.

The present invention even further provides a method of producing a gene product within the cells of a transgenic plant utilising the DNA sequence as described in any one of (a) to (f) above operably associated with the coding sequence of a gene.

The present invention also provides an oligonucleotide probe which selectively hybridizes to a DNA sequence as described in any one of (a) to (f) above.

3.1 SEQUENCE IDENTIFIERS

In the sequence listing:

SEQ ID No. 1 shows the DNA sequence of the isolated barley limit dextrinase inhibitor promoter .

SEQ ID No. 2 shows the nucleotide and derived amino acid sequence of the isolated barley limit dextrinase inhibitor protein.

SEQ ID No. 3 shows the derived amino acid sequence of the isolated barley limit dextrinase inhibitor protein.

SEQ ID No. 4 shows the nucleotide sequence of the PCR primer inhib5.

SEQ ID No. 5 shows the nucleotide sequence of the PCR primer inhib6.

SEQ ID No. 6 shows the nucleotide sequence of the PCR primer AP1.

SEQ ID No. 7 shows the nucleotide sequence of the PCR primer GSP1.

SEQ ID No. 8 shows the nucleotide sequence of the PCR primer AP2

SEQ ID No. 9 shows the nucleotide sequence of the PCR primer GW1.

4. BRIEF DESCRIPTION OF THE FIGURES

In order that the invention may be readily carried into effect reference will now be made, by way of example, to the following diagrammatic drawings in which;

Figure 1 shows the results of an RT-PCR analysis of barley limit dextrinase inhibitor expression. The RT-PCR analysis of LDI expression was performed using Inhib 5 and Inhib 6 primers. Legend: λ/Hind III molecular weight markers (λ); complete barley grains 20 dpa (1); complete barley grains 40 dpa (2); endosperm of barley grains 14 dpa (3); aleurone layers of barley grains treated with gibberellic acid (4); negative control of RT-PCR (5); days post anthesis (dpa);

Figure 2 shows a the results of Northern blot analysis of limit dextrinase expression. The top panel of the blot shows the RNA gel and the bottom panel of the blot shows the Northern blot. Northern blot analysis of LDI expression was performed using LDI RNA probe on different RNA samples. Legend: Leaves in the light (LL); leaves in the dark (LD); roots (r); post anthesis (p.a.);

Figure 3 shows the results of Southern blot analysis of limit dextrinase inhibitor in barley genomic DNA. The Southern analysis of barley ("GoldenPromise") DNA was performed using limit dextrinase inhibitor DNA as a probe. Legend: λ/Hind III molecular weight markers (λ); DNA digested with EcoR V (E); DNA digested with BamH1 (B); DNA digested with Xho I (X);

Figure 4 shows an isolation of limit dextrinase inhibitor promoter fragments by genome walking PCR. DNA was digested with DraI, SmaI, EcoR V, Pvu II and Sca I. The primers used were AP2 and GW 1, and the PCR products analyzed on a preparative 1% TAE agarose gel. Legend: λ/Hind III molecular weight markers (λ); negative control of PCR reaction (-ve). Base pairs (bp);

Figure 5 shows a Barley transformation construct containing the limit dextrinase inhibitor promoter pCAMBIA1302 with LDI promoter;

Figure 6 shows a transient expression assay using the limit dextrinase promoter: GFP construct. Fluorescent microscopy was performed on endosperm 2-4 weeks post anthesis transiently assayed with an LDI Promoter: GFP construct. Legend: Non-bombarded controls (A, C, E and G); bombarded with LDI Promoter:GFP construct and analyzed for GFP expression 24h and 48h after bombardment (B, D, F and H). Scaling bar is 100μm in length.

Figure 7 shows the putative transcription start sites and TATA sequences (-25 bp) within SEQ. ID. No: 1 as predicted by neural network (M.G. Reese, N.L. Harris & F.H. Eeckman. 1996; Large Scale Sequencing Specific Neural Network for promoter & slice site recognition: Biocomputing Proceedings of 1996 Pacific Symposium). The putative transcript start site is shown in larger font.

Figure 8 shows the results of a full sequence search using the BLASTN algorithm for sequences similar to SEQ. ID. No. 1 of the present invention. The search results provided no significant homology with any known sequence in the four databases searched.

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to plant gene regulatory sequences. Specifically this invention relates to a promoter identified in *Hordeum vulgare* (barley).

A plant DNA sequence may be recovered from the cells of the natural host or it may be synthesized directly *in vitro*. Extraction from the natural host enables the isolation *de novo* of novel sequences, whereas *in vitro* DNA synthesis generally requires pre-existing sequence information. Direct chemical *in vitro* synthesis can be achieved by sequential manual synthesis or by automated procedures. DNA sequences may also be constructed by standard techniques of annealing and ligating fragments or by other methods known in the art. Examples of such cloning procedures are given in Sambrook et al.

The DNA sequence of the present invention may be isolated by direct cloning of segments of plant genomic DNA and screening the DNA sequence for the presence of diagnostic sequence motifs characteristic of known promoter sequences. Suitable segments of genomic DNA may be obtained by fragmentation using restriction endonucleases, sonication, physical shearing, or other methods known in the art.

The identification of the cloned segment as a promoter sequence may alternatively be achieved by assessing functionality, for example by linking the cloned segment with a coding sequence derived from a reporter gene and introducing the chimaeric construct into a host cell or cell-free system wherein expression of the reporter gene can be evaluated. This process may form part of another sequence isolation strategy termed promoter trapping, wherein genomic

DNA fragments are cloned directly into "expression vectors" comprising a reporter gene coding region and other sequences necessary for expression in a host cell or cell-free system. The expression may or may not require integration of the chimaeric construct into the host's chromosomal DNA (Topping J F. *et al.* (1991) and Topping J F. *et al.* (1994)).

An alternative method of obtaining DNA sequences of the present invention is by the identification and isolation of a DNA coding sequence which is known to be expressed and subsequently using this sequence to obtain the contiguous promoter sequence, which is by definition directing the expression of the coding sequence. A coding sequence may be obtained by the isolation of messenger RNA (mRNA or polyA+ RNA) from plant tissue. The tissue used for RNA isolation is selected on the basis that suitable gene coding sequences are believed to be expressed in that tissue at optimal levels for isolation.

Various methods for isolating mRNA from plant tissue are well known to those skilled in the art, including for example using an oligo-dT oligonucleotide immobilised on an inert matrix. The isolated mRNA may be used to produce its complementary DNA sequence (cDNA) by use of the enzyme reverse transcriptase (RT) or other enzymes having reverse trancriptase activity. Isolation of an individual cDNA sequence from a pool of cDNAs may be achieved by cloning into bacterial or viral vectors, or by employing the polymerase chain reaction (PCR) with selected oligonucleotide primers. The production and isolation of a specific cDNA from mRNA may be achieved by a combination of the reverse transcription and PCR steps in a process known as RT-PCR.

In order to isolate the promoter sequence of the present invention, a cDNA sequence to a Limit dextrinase inhibitor protein may be identified. Examples of other Limit dextrinase inhibitor protein genes are given in Genbank Accession No.s X13443 for the *Hordeum vulgare* alpha-amylase/trypsin inhibitor; X99982. for the *Triticum aestivum* cDNA for the PUP88 protein ; AJ222975 for the *Hordeum sponteneum* cDNA for the Itr1 gene for BTI-Cme2.2 protein; AP005197 for the *Oryza sativa* cDNA for the putative hageman factor inhibitor protein ; X61032 for the *Triticum durum* cDNA for the alpha amylase inhibitor protein; and X54064 for the *Zea mays* cDNA for the Hageman factor inhibitor protein .

Limit dextrinase inhibitor proteins suitable for use in isolating the promoter of the present invention include proteins having the amino acid sequence given as SEQ. ID. No:3 and proteins homologous to, and having essentially the same biological properties as, the protein disclosed herein as SEQ. ID. No. 3. This definition is intended to encompass natural allelic variations in the protein plus any non-allelic examples. It will be appreciated that the amino acid sequence need not be identical to that of SEQ. ID. No. 3; for the purposes of this invention, the amino acid

sequence may be at least 80%, 90%, or 95% homologous or more with the protein of SEQ. ID. No:3 to retain its biological activity. General categories of potentially equivalent amino acids include, but are not limited to: glutamic acid and aspartic acid; lysine, arginine and histidine; alanine, valine, leucine and isoleucine; asparagine and glutamine; threonine and serine; phenylalanine, tyrosine and tryptophan; and glycine and alanine.

The selected cDNA may then be used to evaluate the genomic features of its gene of origin by use as a hybridisation probe in a Southern blot of plant genomic DNA to reveal the complexity of the genome with respect to that sequence. Alternatively, sequence information from the cDNA may be used to devise oligonucleotides and these may be used in the same way as hybridisation probes for PCR primers to produce hybridisation probes or for PCR primers to be used in direct genome analysis.

Similarly the selected cDNA may be used to evaluate the expression profile of its gene of origin by use as a hybridisation probe in a Northern blot of RNA extracted from various plant tissues, or from a developmental or temporal series. Again sequence information from the cDNA may be used to devise oligonucleotides which can be used as hybridisation probes to produce hybridisation probes or directly for RT-PCR.

The selected cDNA, or derived oligonucleotides, may then be used as a hybridisation probe to challenge a library of cloned genomic DNA fragments and identify overlapping DNA sequences. By this means a contiguous promoter may be identified and isolated.

By the nature of the method of isolation, an isolated cDNA usually comprises the 3' terminus of the coding region and extends towards the 5' terminus. It may not comprise the full-length coding sequence. It is preferable to ensure that the 5' terminal sequence is present if the cDNA is to be used to identify the contiguous promoter. This may be achieved by extension of the cloned cDNA sequence in the 5' direction by a process termed 5' RACE.

If sequence analysis of the cloned cDNA identifies a homologous sequence already reported in the scientific literature, this information may provide a suitable candidate sequence for the 5' terminus. However the possibility of there being different members of the same gene family with similar coding regions, but differing intron regions, promoter sequences and expression profiles may lead to the selection of an incorrect and unsuitable promoter sequence.

Once the 5' terminus of the coding sequence has been identified, the contiguous upstream region containing the promoter may be isolated by further extension in the 5' direction. This may be achieved by methods including vector-ligation PCR, genome walking, vectorette PCR, and other methods. If necessary the process may be repeated with a new primer complementary

to the 5' terminus of the first promoter fragment to ensure that all the control sequences of the promoters are isolated.

The present invention provides a DNA sequence which controls the expression of a Limit dextrinase inhibitor protein which DNA sequence is known herein as SEQ. ID. No. 1, and which may be used as a plant promoter. The present invention also includes within its scope DNA sequences which control the expression of other Limit dextrinase inhibitor protein genes and which are homologous with the DNA sequence of SEQ. ID. No. 1.

Homology may be determined on the basis of percentage identity between two DNA (or polypeptide) sequences. "Percentage identity", as known in the art, is a measure of the relationship between two polynucleotides or two polypeptides, as determined by comparing their sequences. In general the two sequences to be compared are aligned to give a maximum correlation between the sequences. The alignment of the two sequences is examined and the number of positions giving an exact nucleotide (or amino acid) correspondence between the two sequences determined, divided by the total length of the alignment multiplied by 100 to give a percentage identity figure. This percentage identity figure may be determined over the whole length of the sequences to be compared, which is particularly suitable for sequences of the same or very similar lengths and which are highly homologous, or over shorter defined lengths, which is more suitable for sequences of unequal length or which have a lower level of homology.

For example, sequences can be aligned with the software clustalw under Unix which generates a file with the ".aln" extension, this file can then be imported into the Bioedit program (Hall, T.A. 1999. Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98) which opens the .aln file. In the Bioedit window, one can choose individual sequences (two at a time) and align them. This method allows for comparison of entire sequences.

Methods for comparing the identity of two or more sequences are well known in the art. Thus for instance, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux J. et al, Nucleic Acids Res. 12 387-395, 1984, available from Genetics Computer Group, Madison, Wisconsin, USA), for example the programs BESTFIT and GAP, may be used to determine the percentage identity between two polynucleotides and the percentage identity between two polypeptide sequences. BESTFIT uses the "local homology" algorithm of Smith and Waterman (Advances in Applied Mathematics, 2 482-489, 1981) and finds the best single region of similarity between two sequences. BESTFIT is more suited to comparing two polynucleotide or two polypeptide sequences which are dissimilar in length, the program assuming that the shorter sequence represents a portion of the longer. In comparison, GAP

aligns two sequences finding a "maximum similarity" according to the algorithm of Needleman and Wunsch (J. Molec. Biol. 48 443-454, 1970). GAP is more suited to comparing sequences which are approximately the same length and an alignment is expected over the entire length. Preferably the parameters "Gap Weight" and "Length Weight" used in each program are 50 and 3 for polynucleotide sequences and 12 and 4 for polypeptide sequences, respectively. Preferably, percentage identities and similarities are determined when the two sequences being compared are optimally aligned.

Other programs for determining identity and/or similarity between sequences are also known in the art, for instance the BLAST family of programs (Karlin & Altschul, 1990, *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin & Altschul, 1993, *Proc. Natl. Acad. Sci. USA* 90:5873-5877, available from the National Center for Biotechnology Information (NCB), Bethesda, Maryland, USA and accessible through the home page of the NCBI at www.ncbi.nlm.nih.gov). These programs exemplify a preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al., 1990, *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al., 1997, *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules (*Id.*). When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. *See* http://www.ncbi.nlm.nih.gov. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, 1988, *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

Another non-limiting example of a program for determining identity and/or similarity between sequences known in the art is FASTA (Pearson W.R. and Lipman D.J., Proc. Nat. Acac. Sci., USA, 85:2444-2448, 1988, available as part of the Wisconsin Sequence Analysis Package). Preferably the BLOSUM62 amino acid substitution matrix (Henikoff S. and Henikoff J.G., Proc.

Nat. Acad. Sci., USA, 89:10915-10919, 1992) is used in polypeptide sequence comparisons including where nucleotide sequences are first translated into amino acid sequences before comparison.

Yet another non-limiting example of a program known in the art for determining identity and/or similarity between amino acid sequences is SeqWeb Software (a web-based interface to the GCG Wisconsin Package: Gap program) which is utilized with the default algorithm and parameter settings of the program: blosum62, gap weight 8, length weight 2.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, typically exact matches are counted.

Preferably the program BESTFIT is used to determine the % identity of a query polynucleotide or a polypeptide sequence with respect to a polynucleotide or a polypeptide sequence of the present invention, the query and the reference sequence being optimally aligned and the parameters of the program set at the default value.

In the context of the present invention "substantially homologous" sequences are those which have at least 45% sequence identity, preferably at least 50% sequence identity, more preferably at least 60%, 70%, 80%, or 90% sequence identity and most preferably at least 95% sequence identity with the described sequences. In some cases sequence identity may be 98% or more preferably 99%, or more.

The present invention also includes DNA sequences which hybridize to the above isolated DNA, including partial sequences and complementary sequences. Conditions under which such sequences will so hybridize can be determined in a routine manner. For example, the present invention includes within its scope, DNA sequences which hybridise to the DNA sequences of the present invention under stringent conditions. In the context of the present invention, "stringent conditions" are defined as those given in Martin *et al* (EMBO J 4:1625-1630 (1985)) and Davies *et al* (Methods in Molecular Biology Vol 28: Protocols for nucleic acid analysis by non-radioactive probes, Isaac, P.G. (ed) pp 9-15, Humana Press Inc., Totowa N.J, USA)). The conditions under which hybridization and/or washing can be carried out can range from 42°C to 68°C and the washing buffer can comprise from 0.1 x SSC, 0.5 % SDS to 6 x SSC, 0.5 % SDS. Typically, hybridization can be carried out overnight at 65°C (high stringency conditions), 60°C (medium stringency conditions), or 55°C (low stringency conditions). The filters can be washed for 2 x 15 minutes with 0.1 x SSC, 0.5 % SDS at 65°C (high stringency washing). The filters were washed for 2 x 15 minutes with 0.1 x SSC, 0.5 %

SDS at 63°C (medium stringency washing). For low stringency washing, the filters were washed at 60°C for 2 x 15 minutes at 2 x SSC, 0.5% SDS.

The present invention also includes DNA sequences which hybridize to oligonucleotide probes. Preferably the DNA sequences hybridize to oligonucleotide probes under stringent conditions. In instances wherein the nucleic acid molecules are oligonucleotides ("oligos"), highly stringent conditions may refer, for example, to washing in 6x SSC / 0.05% sodium pyrophosphate at 37°C (for 14 base oligos), 48°C (for 17 base oligos), 55°C (for 20 base oligos), and 60°C (for 23 base oligos).

The gene coding sequence employed in carrying out the present invention may be active in some or all plant tissues. The sequence employed may encode a protein or an RNA moiety. Through recombinant DNA techniques the sequence may encode a synthetic variant of a protein or RNA, a partial sequence, or a composite sequence comprising regions from one or more genes. For example, the chimaeric gene may encode a polyprotein. The sequence may also comprise repeated, inverted or complementary sequences, to achieve disruption of the transcription and translation of an endogenous gene or genes by, for example, antisense or RNA inhibition technology.

The sequence known herein as SEQ. ID. No. 2 includes a coding sequence therein for the gene encoding a limit dextrinase inhibitor protein. In addition, many plant, bacterial and viral genes may be actively expressed. Exemplary genes encode the GUS, GFP and luciferase enzymes which may be used as reporter genes for promoter function. In addition the LDI promoter may be used to direct the expression of DNA sequences designed to alter the metabolism of the endosperm of the seeds of cereal plants, especially barley. Such genes include those for carbohydrate metabolism, starch metabolism, amino acid and protein metabolism and lipid metabolism.

The gene product of the present invention can be produced by recombinant techniques, wherein genomic DNA clones or cDNA clones for the DNA coding sequence are produced, isolated, proliferated, and incorporated into a plant transformation vector of the present invention.

Plant transformation vectors

Plant transformation vectors of the present invention will contain "expression cassettes" comprising 5'-3' in the direction of transcription, a promoter sequence as described in the present invention, a gene coding sequence as discussed above and, optionally a 3' untranslated,

terminator sequence including a stop signal for RNA polymerase and a polyadenylation signal for polyadenylase.

The promoter sequence may be present in one or more copies, and such copies may be identical or variants of the promoter sequence as described above. Such copies may also be complete or partial sequences as described above.

The terminator sequence may be obtained from plant, bacterial or viral genes. Suitable terminator sequences are the pea *rbcS* E9 terminator sequence, the *nos* terminator sequence derived from the nopaline synthase gene of *Agrobacterium tumefaciens* and the 35S terminator sequence from cauliflower mosaic virus, for example. A person skilled in the art will be readily aware of other suitable terminator sequences.

The expression cassette may also comprise a gene expression enhancing mechanism to increase the strength of the promoter. An example of such an enhancer element is that derived from a portion of the promoter of the pea plastocyanin gene, and which is the subject of International Patent Application, Publication No. WO 97/20056.

These regulatory regions may be derived from the same gene as the promoter DNA sequence of the present invention or may be derived from different genes, e.g. from *Hordeum vulgare* or other organisms. All of the regulatory regions should be capable of operating in cells of the tissue to be transformed.

The gene coding sequence may be derived from the same gene as the promoter DNA sequence of the present invention or may be derived from a different gene, e.g. from *Hordeum vulgare* or another organism.

The expression cassette may be incorporated into a basic plant transformation vector, such as, binary or super-binary vectors, for example, pBIN19 and pBIN+ and other suitable plant transformation vectors known in the art.

In addition to the expression cassette, the plant transformation vector will contain such sequences as are necessary for the transformation process. These may include the *Agrobacterium vir* genes, one or more T-DNA border sequences, and a selectable marker or other means of identifying transgenic plant cells.

Production of the gene product

Expression of a DNA coding sequence in the plant host cell will produce an RNA transcript. If the coding sequence is derived from a structural gene, the RNA transcript is then translated into a protein gene product. If desired, the gene product can be isolated by standard techniques for isolating proteins from biological systems, such as salt precipitation, column

chromatography, immunoaffinity techniques, electrophoresis, recrystallisation, centrifugation, and such like.

Example 1. Isolation of barley *limit dextrinase inhibitor* cDNA fragments.

Cloning of LDI gene by RT-PCR

Total RNA was extracted from barley (var. Golden Promise) grains 2 and 4 weeks post anthesis with a LiCl method as described by (Cathala *et al.*, 1983).

2 µg of RNA was treated with Rnase-free DNase I (Amersham Pharmacia Biotech) and used to synthesize 30 µl first strand cDNA using random hexamer primers (Roche) and M-MLV reverse transcriptase (Promega) using the reaction conditions recommended by the manufacturer.

It will be recognised by one skilled in the art that other mRNA extraction and cDNA synthesis methods exist which could be employed to produce cDNA from tissue of *Hordeum vulgare.*

A 3 µl aliquot of the cDNA product was used in a standard PCR reaction containing 2 mM $MgCl_2$, 8% (v/v) DMSO and primers Inhib-5 (5'-ACCAATAAACTAGTATCAACAATGGCATCCGACCA-3' SEQ ID No 4) and Inhib-6 (5'-CCAACCTTTTTTATTCATCAATCGGCCACA-3' SEQ ID No 5), which were designed against the *Hordeum vulgare* Limit dextrinase inhibitor sequence (Genbank Accession no. X13443.), using Taq polymerase (Bioline) as recommended by the manufacturer. The PCR program used was 5 min at 94 °C, followed by 35 cycles of 94 °C for 30 sec, 63.5 °C for 30 sec and 72 °C for 1 min, finalized by 7 min at 72 °C. The product length was 623 bp. The amplified product was cloned into a pGEM-T Easy vector (Promega) and verified by sequencing.

It will be recognised by one skilled in the art that other methods exist which could be employed to produce cloned DNA fragments.

The sequence of this cDNA clone is shown in SEQ ID No 2. Compared to the published sequence (Genbank Accession no. X13443.), there are five single base pair substitutions, which in turn lead to two amino acid substitutions.

Example 2. RT-PCR analysis of Limit Dextrinase Inhibitor expression.

Total RNA was isolated from barley tissues using a LiCl method (M.-J.Cho *et al.*, 1998). 2 µg of total DNAse treated RNA was used to produce cDNA with M-MLV reverse transcriptase. 1/10 of this cDNA was used to run a PCR reaction with specific primers for the entire coding region of LDI.

The expression of the LDI gene was investigated by using RT-PCR with the specific primers Inhib 5 (SEQ ID No. 4) and Inhib 6 (SEQ ID No. 5) on RNA from developing grains. Expression of LDI was found in grains 20 and 40 days post anthesis, the endosperm and gibberellic acid treated aleurone layers. The highest level of expression was found in the endosperm. As shown in Figure 1.

Example 3. Northern analysis of Limit Dextrinase Inhibitor expression.

Total RNA was isolated from developing and germinating barley grains, barley leaves in the light or dark, and barley roots using a LiCl method (M.-J.Cho et al., 1998). 5 µg of total RNA were separated on a 1 % agarose gel with formaldehyde. The RNA was blotted over night with 10× SSC onto a positively charged Nylon membrane and UV crosslinked. The blot was then prehybridized in DIG-EasyHyb and hybridised with 100 ng/mlDIG-labeled RNA probe at 68 °C over night. The blot was washed once in 2× SSC, followed by 2 washes with 0.2× SSC at 68 °C. After that the blot was blocked for 1 hour at RT in blocking reagent. This was followed by 30 minutes in 1:10,000 dilution of anti-DIG antibody solution. To remove excess antibody the blot was washed 4× in Tris buffer with salt, then equilibrated in phosphatase buffer and the chemiluminiscent substrate (CDP-star) added. Chemiluminescence detection using X-ray film was performed. Northern blotting and hybridisation was carried out as detailed by manufacturers of the DIG label (Roche).

Expression of LDI was only detected in early development of barley grains (2 and 4 weeks post anthesis). No expression was found in germinating grains and vegetative tissue as shown in Figure 2.

Example 4. Southern blotting

Genomic DNA of barley leaf tissue was isolated according to the method of A.W.MacGregor *et al.*, 1995 and S.L.Delaporta *et al.*, 1983. 10 µg digested genomic DNA (isolated using the Ambion Phytopure Kit according to manufacturer's instructions) was separated on a 0.8 % TAE agarose gel at 20V over night. The DNA was depurinated in 250 mM HCl for 10 minutes, denatured for 30 min in 0.4 M NaOH and neutralized before blotting for 8 hours onto a positively charged Nylon membrane (Roche) in 20x SSC. The DNA was UV crosslinked and prehybridized in DIG-Easyhyb for 1 hour at 52 °C and hybridised with 40ng/ml DIG labeled DNA probe at 52 °C over night. Washing and detection was carried out as described before for Northern blotting except for a more stringent wash with 0.1 x SSC.

Southern blot analysis was used to examine the number of gene copies in wildtype (wt) "Golden Promise" barley. The exact number of copies of LDI in wildtype "Golden Promise" barley is hard to estimate as several bands per digest are visible. The most likely explanation is that there is a small gene family of related genes present in barley, as shown in Figure 3.

Example 5. Isolation of the limit dextrinase inhibitor promoter region by genome walking.

Genomic DNA of barley leaf tissue was isolated according to the method of A.W.MacGregor etal., 1995 and S.L.Delaporta et al., 1983. 2.5 µg genomic barley DNA (variety "Golden Promise") was digested with the restriction enzymes Sca I, Pvu II, EcoR V, Sma I or Dra I respectively. These DNA fragments were then ligated to an asymmetric adapter and used as a template for PCR reactions with adapter primers and gene specific primers to isolate the 5' region of the LDI gene. All procedures were carried out as described in the Clontech manual. The first PCR was carried out using primers AP1 (SEQ ID No. 6) and GSP1 (SEQ ID No. 7). 1µl of this 50 µl PCR reaction was used as template for the secondary (nested) PCR using primers AP2 (SEQ ID No. 8) and GW1 (SEQ ID No. 9).

Using the procedures described above, two candidate fragments were isolated, from digests with Sma I and Pvu II, as shown in Figure 4. The DNA fragments from the Sma I (~1000 bp) and Pvu II (~ 650 bp) digests were purified and cloned into the Topo 4 vector and sequenced. Both fragments (Sma I and Pvu II) showed the same A and T rich sequence, spanning the 790 bp long 5' region of LDI. The promoter sequence is shown in SEQ ID No. 1. SEQ. ID. No. 1 discloses a nucleic acid sequence of 833 bp, within which sequence there is the 790 bp promoter sequence of the present invention. The sequence also contains therewithin the 5' untranslated region of the gene coding sequence, up to and including the first codon of the gene.

Example 6. Transient Expression of Limit Dextrinase Inhibitor Promoter marker gene constructs.

In order to investigate if the 790 bp region upstream of the LDI gene was sufficient as a promoter, a vector containing the putative promoter region plus a reporter gene was constructed. It consists of the 790 bp *Sma* I fragment, the GFP gene (green fluorescent protein) and a terminator (NOS). This construct (Figure 5) was used for transient assays on barley endosperm (2-4 weeks post anthesis) and for stable barley transformation. Endosperms were transformed with a biolistic device using .construct DNA coated gold particles and analysed by fluorescence microscopy after 24 and 48 hours post bombardment. The results in shown in Figure 6

demonstrate the expression of GFP, showing that the isolated genomic sequence acts as promoter.

Example 7. Stable expression of Limit Dextrinase Inhibitor Promoter marker gene constructs.

Stable barley transformation of immature embryos of the barley variety "Golden Promise" was performed with a biolistic device using DNA coated gold particles. Selection and regeneration procedures were performed as described by Y.Wan & P.G.Lemaux, 1994.

References.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25** 3389-3402

Ausubel, F.M., Bent, R., Kingston, R.E., Moore, D.D., Seidmann, J.G., Smith, J.A. & Struhl, K. (1989). *Current protocols in molecular biology* . New York: Greene and Willy Interscience.

Benfey, P.N., and Chua, N-H. (1989) Regulated genes in transgenic plants. Science **244**, 174-181.

Bevan, M. (1984) Binary *Agrobacterium* Vectors for Plant Transformation. Nucleic Acids Research **12** 8711-8721

Buluwela, L., Forster, A., Boehm, T. & Rabbitts, T.H. (1989). A rapid procedure for colony screening using nylon filters. *Nucleic Acids Research* **17**, 452-452.

Christensen, A. H. & Quail, P.H. (1996). Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research* **5**, 213-218

Cho, M.-J., Jiang, W. & Lemaux, P.G. (1998). Transformation of recalcitrant barley cultivars through improvement of regenerability and decreased albinism. *Plant Science* **138**, 229-244.

Conkling, M.A., Cheng, C-L., Yamamoto, Y.T. and Goodman, H.M. (1990) Plant Physiol. **93** 1203-1211

Delaporta, S.L., Wood, J. R. & Hicks, J.B. (1983). A plant DNA minipreparation: version 2.*Plant Molecular Biology Reporter* **1**, 19-22.

Gatz, C. (1995) Novel inducible/repressible gene expression systems. Methods in Cell Biol. **50**, 411-424

Horsch, R.B., Fry, J.E., Hoffman, N.L., Eichholtz, D., Rogers, S.G. & Fraley, R.T. (1985) A simple and general method for transferring genes into plants. Science **227**, 1229-1231.

Jefferson, R.A. (1987) Assaying Chimeric Genes in Plants: The GUS Gene Fusion System. Plant Molec. Biol. Reptr. **5** 387-405

Kristensen, M., Planchot, V., Abe, J.-I. & Svensson, B (1998). Large-scale purification and characterization of barley limit dextrinase, a member of the a-amylase structural family. *Cereal Chemistry* **75**, 473-479.

Lazaro, A., Rodriguez-Palenzuela, P., Marana, C., Carbonero, P. & Garciaolmedo, F. (1988). Signal peptide homology between the sweet protein thaumatin II and unrelated cereal alpha-amylase/trypsin inhibitors. *FEB* **239**, 147-150.

Lemaux, P.G. Cho, M.-J., Zhang, S. & Bregitzer, P. (1999). Transgenic cereals: Hordeum vulgare L. (barley). In *Molecular improvement of cereal crops*, ed. I.K.VASIL, pp. 255-316. Dordrecht: Kluwer Academic Publishers.

Macgregor, A.W., Macri, L.J., Bazin, S.L. & Sadler, G.W. (1995). Limit dextrinase inhibitor in barley and malt and its possible role in malting and brewing. *EBC Congress* 185-193.

Macgregor, A.W., Macri, L.J., Schroeder, S.W. & Bazin, S.L. (1994). Purification and characterisation of limit dextrinase inhibitors from barley. *Journal of Cereal Science* **20**, 33-41.

Macgregor, E.A., Bazin, S.L., Ens, E.W., Lahnstein, J., Macri, L.J., Shirley N.J. & Macgregor, A.W. (2000). Structural models of limit dextrinase inhibitors from barley. *Journal of Cereal Science* **31**, 79-90.

Manners, D.J. (1997). Observations on the possible involvement of debranching enzymes in starch biosynthesis. *J.Appl.Glycosci.* **44**, 541-546.

Sambrook, J., Frisch, E.F. and Maniatis, T. (Eds.) (1989) Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York.

Topping J F; Wei W; Lindsey K; (1991) Development 112, 1009-1019. Functional tagging of regulatory elements in the plant genome

Topping J F; Agyeman F; Henricot B; Lindsey K (1994) The Plant Journal 5, 895-903, Identification of molecular markers of embryogenesis in Arabidopsis thaliana by promoter trapping.)

Wan, Y. & Lemaux, P.G. (1994). Generation of large numbers of independently transformed fertile barley plants. *Plant Physiology* **104**, 37-48.

SEQUENCE LISTING

<110> Advanced Technologies (Cambridge) Ltd.

<120> Limit Dextrinase Inhibitor Promoter

<130> RD-ATC-33

<140>
<141>

<160> 9

<170> PatentIn Ver. 2.1

<210> 1
<211> 833
<212> DNA
<213> Hordeum vulgare

<400> 1
catcgattac tatagggcac gcgtggtcga cggcccgggc tgttattgga caccaaatgt 60
atcataaact tgtttttca ccgacaaaat attgctcctc catttcgcat taaaattgtc 120
aagcatgctt gcaacagtaa cacgaacatt cataaaaaaa atatttttta agaaaacatt 180
tactattttt ttgttactat tcatctggga gcatgtgctt ccggaagcca aaatgcccct 240
tccaatatgc cccgtgtaaa agaaacccct tctttcctaa aaatatatat catcgtccgt 300
catgatacgt ttatgtattc aacgaaaaat attttcgcat gtcaccaaaa atgttttata 360
ttacacaagt gaacaaatat gataaactcc ctcgtgttaa ctatttttc tgtgaaataa 420
aaggatgaca atcaaaacaa aaatgtagac tgtaaacaaa gaaaacatta tttcctagaa 480
ataaaaaaaa agattagagg gatatgtatt gtcgaaacac atgaggacta gaacaaaaga 540
aaaagggaaa tgagaaggaa aaaaggggta accattaccc aaagaaaaca gaaagtaaac 600
tagacgtgtc gaagggaaac ggagtttgca ggggcgttcc aaattcagtt gcaagaacct 660
ccaaataaac gccaacaaga aagaaatgag cattacttgc gcgctttgca ctcttatctc 720
tagcatctcc cgatacatac atacatgtag cctagctgca gatcttgaat agctattctt 780
gcccaccagg ccaagagatt gaaccaacga ccaataaact agtatcaaca atg 833

<210> 2
<211> 517
<212> DNA
<213> Hordeum vulgare

<220>
<221> CDS
<222> (14)..(457)

<400> 2
actagtatca aca atg gca tcc gac cat cgt cgc ttc gtc ctc tcc ggc 49
Met Ala Ser Asp His Arg Arg Phe Val Leu Ser Gly
1 5 10

gcc gtc ttg ctc tcg gtc ctc gcc gtc gcc gcc gcc acc ctg gag agc 97
Ala Val Leu Leu Ser Val Leu Ala Val Ala Ala Ala Thr Leu Glu Ser
15 20 25

gtc aag gac gag tgc caa cca ggg gtg gac ttc ccg cat aac ccg tta 145
Val Lys Asp Glu Cys Gln Pro Gly Val Asp Phe Pro His Asn Pro Leu
30 35 40

gcc acc tgc cac acc tac gtg ata aaa cgg gtc tgc ggc cgc ggt ccc 193
Ala Thr Cys His Thr Tyr Val Ile Lys Arg Val Cys Gly Arg Gly Pro
45 50 55 60

```
agc cgg ccc atg ctg gtg aag gag cgg tgc tgc cgg gag ctg gcg gcc   241
Ser Arg Pro Met Leu Val Lys Glu Arg Cys Cys Arg Glu Leu Ala Ala
                65                  70                  75

gtc ccg gat cac tgc cgg tgc gag gcg ctg cgc atc ctc atg gac ggg   289
Val Pro Asp His Cys Arg Cys Glu Ala Leu Arg Ile Leu Met Asp Gly
            80                  85                  90

gtg cgc acg ccg gag ggc cgc gtg gtt gag gga cgg ctc ggt gac agg   337
Val Arg Thr Pro Glu Gly Arg Val Val Glu Gly Arg Leu Gly Asp Arg
        95                  100                 105

cgt gac tgc ccg agg gag gag cag agg gcg ttc gcc gcc acg ctt gtc   385
Arg Asp Cys Pro Arg Glu Glu Gln Arg Ala Phe Ala Ala Thr Leu Val
    110                 115                 120

acg gcg gcg gag tgc aac cta tcg tcc gtc cag gag ccg gga gta cgc   433
Thr Ala Ala Glu Cys Asn Leu Ser Ser Val Gln Glu Pro Gly Val Arg
125                 130                 135                 140

ttg gtg cta ctg gca gat gga tga cgatcgaaat gcgccaaggt aatgaagcgg  487
Leu Val Leu Leu Ala Asp Gly
                145

agtactgtat acagaataaa agtactcgag                                  517
```

<210> 3
<211> 147
<212> PRT
<213> Hordeum vulgare

<400> 3

```
Met Ala Ser Asp His Arg Arg Phe Val Leu Ser Gly Ala Val Leu Leu
  1                 5                  10                  15
Ser Val Leu Ala Val Ala Ala Ala Thr Leu Glu Ser Val Lys Asp Glu
            20                  25                  30
Cys Gln Pro Gly Val Asp Phe Pro His Asn Pro Leu Ala Thr Cys His
        35                  40                  45
Thr Tyr Val Ile Lys Arg Val Cys Gly Arg Gly Pro Ser Arg Pro Met
    50                  55                  60
Leu Val Lys Glu Arg Cys Cys Arg Glu Leu Ala Ala Val Pro Asp His
65                  70                  75                  80
Cys Arg Cys Glu Ala Leu Arg Ile Leu Met Asp Gly Val Arg Thr Pro
                85                  90                  95
Glu Gly Arg Val Val Glu Gly Arg Leu Gly Asp Arg Arg Asp Cys Pro
            100                 105                 110
Arg Glu Glu Gln Arg Ala Phe Ala Ala Thr Leu Val Thr Ala Ala Glu
        115                 120                 125
Cys Asn Leu Ser Ser Val Gln Glu Pro Gly Val Arg Leu Val Leu Leu
    130                 135                 140
Ala Asp Gly
145
```

<210> 4
<211> 35
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PCR primer

<400> 4
accaataaac tagtatcaac aatggcatcc gacca 35

<210> 5
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 5
ccaacctttt ttattcatca atcggccac 29

<210> 6
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 6
gtaatacgac tcactatagg gc 22

<210> 7
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 7
agacccgttt tatcacgtag gtgtgg 26

<210> 8
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 8
actatagggc acgcgtggt 19

<210> 9
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR primer

<400> 9
agacggcgcc ggagaggacg aagcga 26

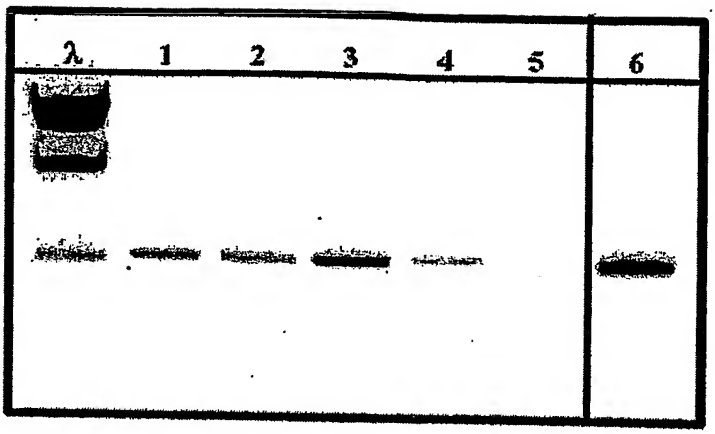

Figure 1. RT-PCR analysis of barley limit dextrinase inhibitor expression.

Figure 2. Northern blot analysis of limit dextrinase expression.

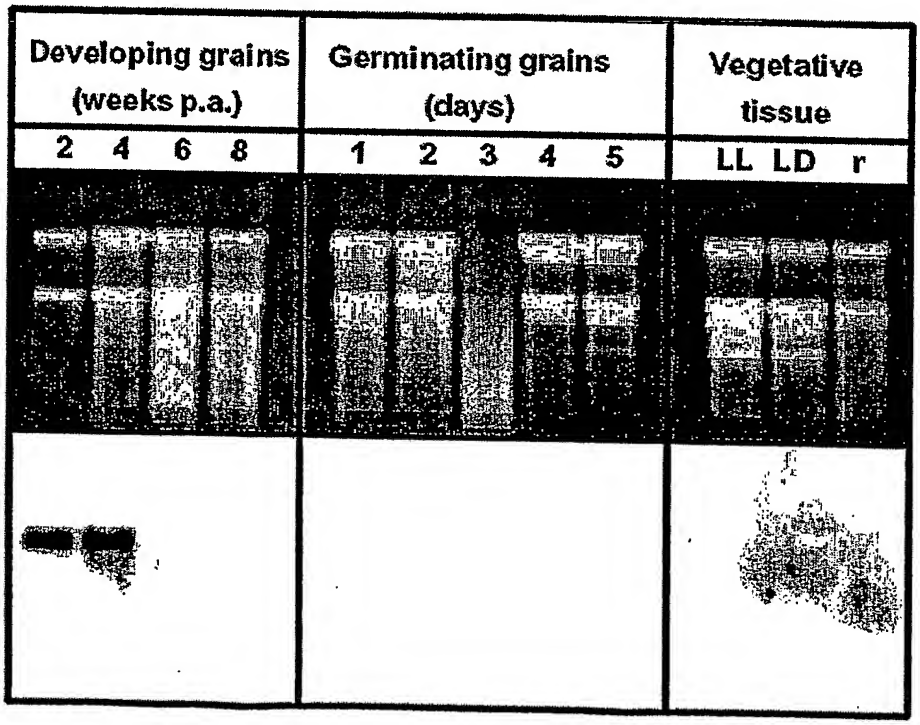

Figure 3. Southern blot analysis of limit dextrinase inhibitor in barley genomic DNA.

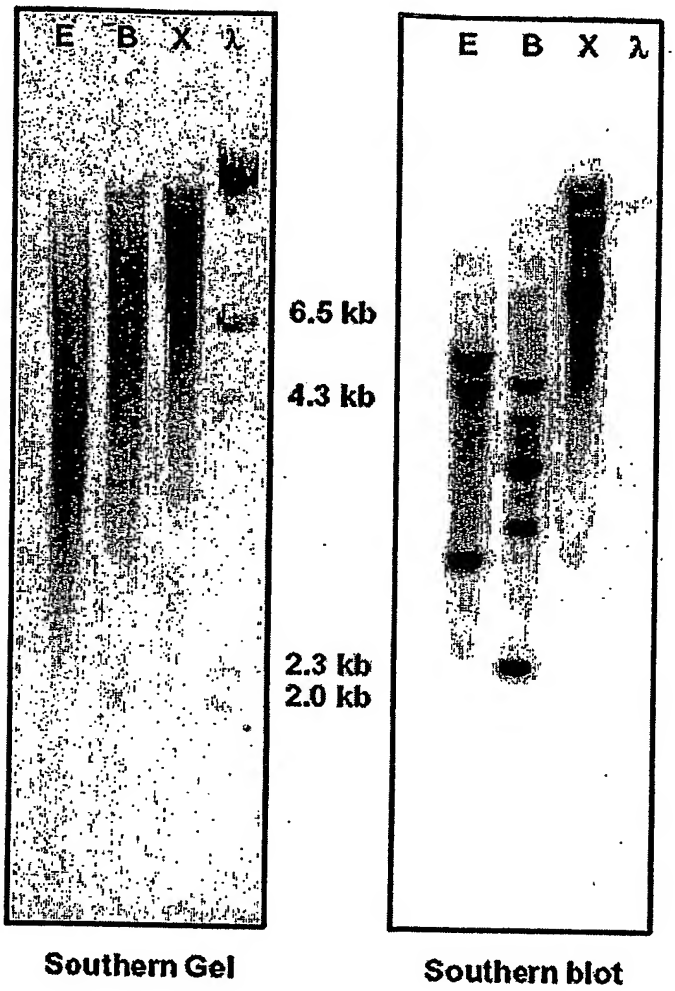

Figure 4. Isolation of limit dextrinase inhibitor promoter fragments by genome walking PCR.

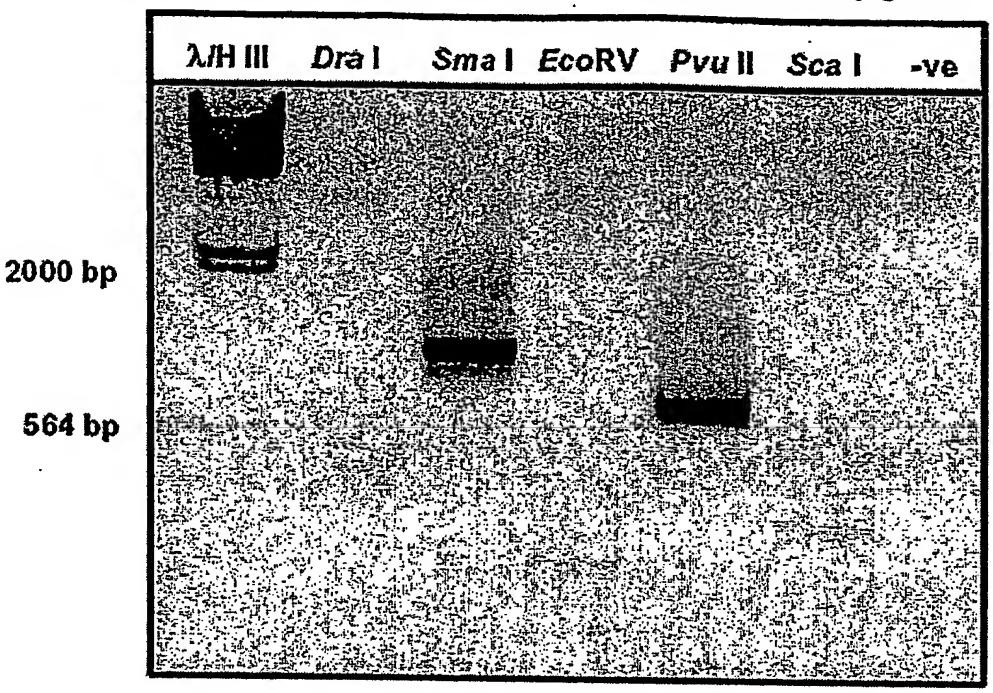

Figure 5. Barley transformation construct containing the limit dextrinase inhibitor promoter.

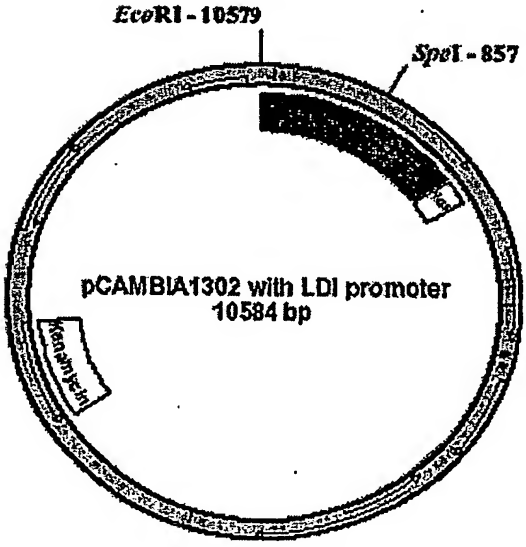

Figure 6. Transient expression assays using the limit dextrinase promoter: GFP construct.

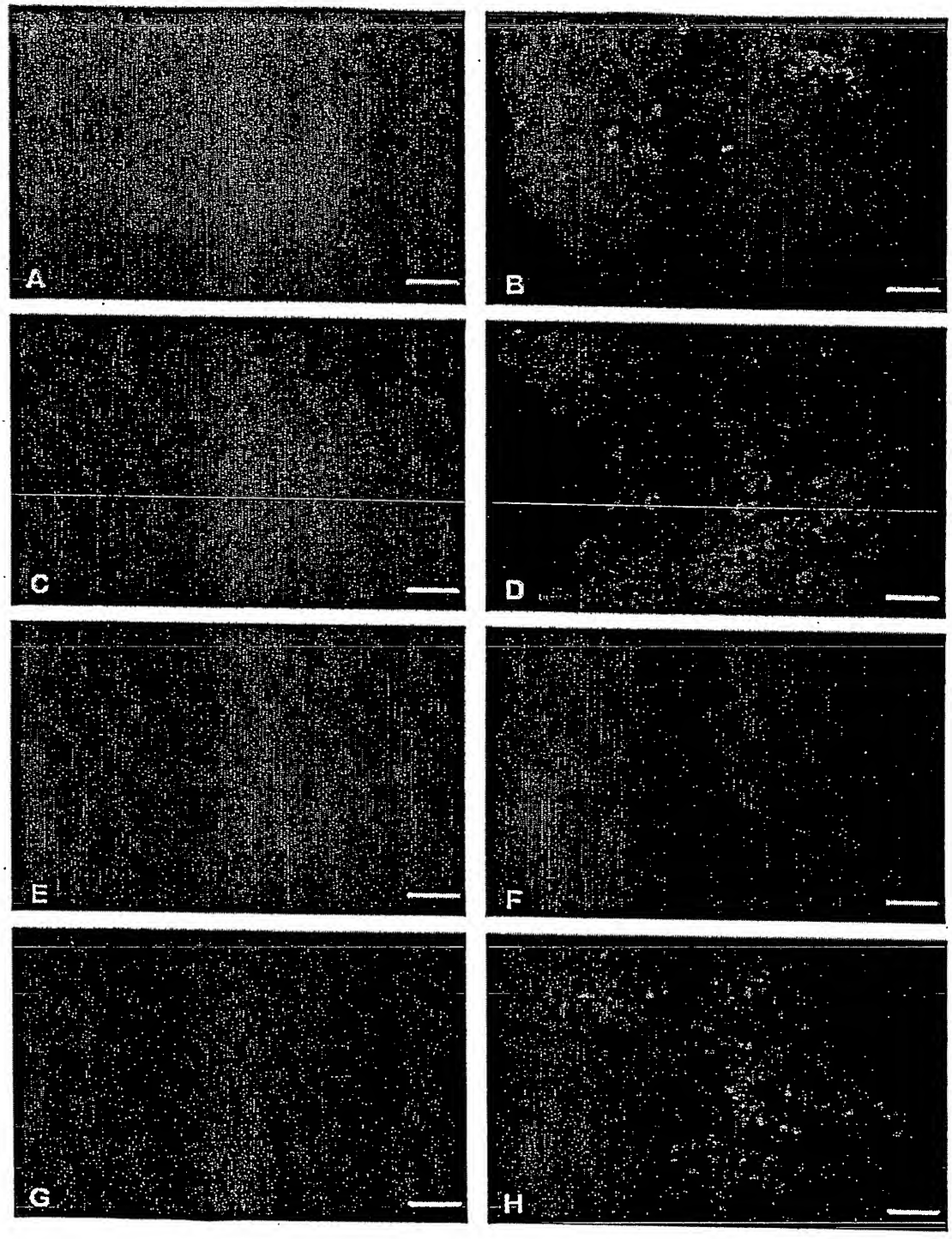

Figure 7.

Start	End	Score	Promoter Sequence
247	297	0.89	ATGCCCCGTGTAAAAGAAACCCCTTCTTTCCTAAAAATATATATCATCGT
273	323	0.95	TTTCCTAAAAATATATATCATCGTCCGTCATGATACGTTTATGTATTCAA
652	702	0.83	CAAGAACCTCCAAATAAACGCCAACAAGAAAGAAATGAGCATTACTTGCG

Figure 8

Sequences producing significant alignments:	Score (bits)	E Value
gi|18868|emb|X13443.1|HVAATI Barley mRNA for alpha-amylase/...	82	2e-12
gi|22035404|gb|AY124482.1| Danio rerio myogenin gene, exons...	80	6e-12
gi|24614584|gb|AY050653.1| Giardia intestinalis SEC24 (SEC2...	76	1e-10
gi|10938038|gb|AF162890.1|AF162890S1 Mus musculus peroxisom...	76	1e-10
gi|5821237|dbj|AB021922.1| Homo sapiens gene for lectin-lik...	76	1e-10
gi|3721562|dbj|AB011276.1| Mus musculus gene for alpha1a ca...	74	4e-10
gi|30313388|gb|AY099112.1| Rattus norvegicus obese protein ...	72	2e-09
gi|18140057|gb|AF457660.1| Castanea dentata clone ACS2 vasc...	72	2e-09
gi|4104807|gb|AF039526.1|AF039526 Homo sapiens MHC class I ...	72	2e-09
gi|8050595|gb|AF233737.1|AF233737 Agrotis ipsilon AiC5 chym...	72	2e-09
gi|6690643|gb|AF191544.1|AF191544 Homo sapiens estrogen rec...	72	2e-09
gi|6164589|gb|AF051769.1|AF051769 Homo sapiens hyaluronidas...	72	2e-09
gi|3643823|gb|AF075270.1|AF075270 Hordeum vulgare high affi...	72	2e-09
gi|19919401|gb|AF435445.1| Pleurotus ostreatus manganese pe...	70	6e-09
gi|8132114|gb|AF153014.1| Trichomonas vaginalis Tvp14 (tvp1...	70	6e-09
gi|4456992|gb|AF077743.1|AF077743 Mus musculus transcriptio...	70	6e-09
gi|12276207|gb|AF269146.1|AF269146 Bilophila wadsworthia ta...	70	6e-09
gi|14275833|emb|AJ289605.1|MMU289605 Mus musculus partial L...	70	6e-09
gi|3378604|emb|AJ009889.1|HSAJ9889 Homo sapiens GHR gene, V...	70	6e-09
gi|3916231|gb|AF074397.1|AF074397 Homo sapiens anti-mulleri...	70	6e-09
gi|5139506|emb|Z18892.2|MMDESMINP Mus musculus desmin gene	70	6e-09
gi|15487305|dbj|AB060299.1| Mus musculus gene for acetyl Co...	70	6e-09
gi|12697590|dbj|AB046716.1| Homo sapiens hST3Gal I gene for...	70	6e-09
gi|25453365|gb|AY050651.2| Giardia intestinalis MYB (MYB) m...	68	2e-08
gi|13183059|gb|AF237414.1| Anaplasma phagocytophilum RNA po...	68	2e-08
gi|13625520|gb|AY014277.1| Lolium perenne gibberellin 20-ox...	68	2e-08
gi|17105179|gb|AF439558.1|AF439558 Mus musculus X2CR1 gene,...	68	2e-08
gi|16209547|gb|AY052528.1| Glycine max calmodulin isoform-4...	68	2e-08
gi|15213480|gb|AF241535.1|AF241535 Homo sapiens mucin 4 (MU...	68	2e-08
gi|15558849|emb|AJ310936.1|FSO310936 Fusarium solani chy ge...	68	2e-08
gi|4878023|gb|AF131239.2|AF131239 Rattus norvegicus alpha 1...	68	2e-08
gi|15216031|emb|AJ318812.1|VFA318812 Vicia faba var. minor ...	68	2e-08
gi|4454294|emb|AJ132779.1|MMU132779 Mus musculus p107 gene ...	68	2e-08
gi|3133088|emb|Y16736.1|HSA16736 Homo sapiens dif-2 gene, p...	68	2e-08
gi|6690521|gb|AF154245.1|AF154245 Rattus norvegicus chemota...	68	2e-08
gi|5731977|gb|AF114032.1|AF114032 Mus musculus glycogenin-1...	68	2e-08
gi|26453412|dbj|AB094665.1| Seriola quinqueradiata YGHL1 ge...	68	2e-08
gi|4039145|gb|AF099083.1|AF099083 Homo sapiens growth hormo...	68	2e-08
gi|2739123|gb|AF029214.1|MMOX2S1 Mus musculus MRC OX-2 anti...	68	2e-08
gi|2895903|gb|AF046916.1|AF046916 Ruminococcus flavefaciens...	68	2e-08
gi|1916583|gb|U53907.1|RNU53907 Rattus norvegicus microsate...	68	2e-08

gi|6683556|dbj|AB024534.1| Rattus norvegicus gene for thiaz...	68	2e-08
gi|7109286|gb|AF227508.1| Rattus norvegicus intestinal alka...	66	1e-07
gi|13345792|gb|AF332759.1| Homo sapiens partially duplicate...	66	1e-07
gi|5002511|emb|Z98266.1|HSZ98266 Homo sapiens gene encoding...	66	1e-07
gi|6560627|gb|AF112228.1|HSCD30P1 Homo sapiens CD30 protein...	66	1e-07
gi|24475537|dbj|AB084484.1| Betula platyphylla DNA, microsa...	66	1e-07
gi|28875405|gb|AF515463.1| Biomphalaria glabrata fibrinogen...	66	1e-07
gi|18140058|gb|AF457661.1| Castanea dentata clone ACS10A va...	64	4e-07
gi|2581766|gb|U77633.1|RNU77633 Rattus norvegicus chromosom...	64	4e-07
gi|2764739|emb|AJ002743.1|HSAJ2743 Homo sapiens cardiotroph...	64	4e-07
gi|7579914|emb|AJ277249.1|HSA277249 Homo sapiens partial HR...	64	4e-07
gi|1518845|gb|U63899.1|MMU63899 Mus musculus Girk2 gene, pr...	64	4e-07
gi|7109284|gb|AF227507.1| Rattus norvegicus intestinal alka...	62	2e-06
gi|12744733|gb|AF318503.1| Danio rerio Myod (myod) gene, co...	60	6e-06
gi|2581767|gb|U77634.1|RNU77634 Rattus norvegicus chromosom...	60	6e-06
gi|14251200|gb|AF220499.2|AF220499 Acidithiobacillus ferroo...	60	6e-06
gi|14043019|gb|AF221946.2|AF221946 Rickettsia rickettsii ce...	60	6e-06
gi|2342636|emb|Y11638.1|MMY11638 M.musculus CYP4A14 gene, e...	60	6e-06
gi|18873678|emb|AJ272507.1|HSA272507 Homo sapiens partial K...	60	6e-06
gi|2564335|dbj|AB008218.1|AB008218S1 Homo sapiens gene for ...	60	6e-06
gi|17907575|emb|AJ409277.1|CDR409277 Camelus dromedarius pa...	58	2e-05
gi|5091690|gb|AF139181.1|AF139181 Bartonella henselae S-ade...	58	2e-05
gi|14164368|dbj|AB052355.1|AB05234S16 Mus musculus gene for...	58	2e-05
gi|15081477|gb|AF401090.1| Wolbachia pipientis RNA polymera...	56	9e-05
gi|17298240|gb|AF283339.1|F283327S13 Homo sapiens candidate...	56	9e-05
gi|4139055|gb|AF072833.1|AF072833 Homo sapiens SP23 gene, p...	56	9e-05
gi|17298258|gb|AF283357.1|F283327S31 Homo sapiens candidate...	54	4e-04
gi|18413572|emb|AJ428930.1|XLA428930 Xenopus laevis partial...	54	4e-04
gi|18873687|emb|AJ272516.1|HSA272516 Homo sapiens partial K...	54	4e-04
gi|29648446|gb|AY190007.1| Pan paniscus clone BoE26-M13R LI...	52	0.001
gi|22651882|gb|AF291761.1| Ipomoea batatas S-adenosylmethio...	52	0.001
gi|22347797|gb|AF532732.1| Danio rerio mx gene, promoter se...	52	0.001
gi|11230634|emb|AJ289159.1|HSA289159 Homo sapiens CD30 gene...	52	0.001
gi|3342093|gb|AF074905.1|HOMOSLC04 Homo sapiens neuronal an...	50	0.006
gi|13377504|gb|AF325198.1|AF325198 Triticum aestivum LRR14 ...	48	0.023
gi|6649909|gb|AF026274.1|AF026274 Mus musculus Sumiko (sumi...	48	0.023
gi|6224791|gb|AF190816.1|AF190816 Homo sapiens complement f...	48	0.023
gi|4104439|gb|AF035664.1|HSTGFRBI3 Homo sapiens transformin...	48	0.023
gi|21213045|emb|AJ487974.1|EAM487974 Enterobacter amnigenus...	48	0.023
gi|15638904|gb|AC024246.8| Homo sapiens BAC clone RP11-643O...	46	0.090
gi|15638822|gb|AC079120.6| Homo sapiens BAC clone RP11-345M...	46	0.090
gi|13793999|gb|AY029002.1| Pinus taeda isolate PTLTP3-iii27...	46	0.090
gi|11093799|gb|AF285184.1|AF285184 Mus musculus basic trans...	46	0.090
gi|2984654|gb|AF039088.1|AF039088 Homo sapiens non-hepatic ...	46	0.090
gi|3342092|gb|AF074904.1|HOMOSLC03 Homo sapiens neuronal an...	44	0.36
gi|29029467|gb|AY173030.1| Danio rerio zinc finger transcri...	44	0.36

gi|2281684|gb|AF009433.1|AH006708S06 Homo sapiens clone 22 ...	44	0.36
gi|13936217|gb|AY029021.1| Pinus radiata isolate PRLTP3-iii...	44	0.36
gi|13936206|gb|AY029014.1| Pinus taeda isolate PTLTP9-iii15...	44	0.36
gi|13936204|gb|AY029013.1| Pinus taeda isolate PTLTP4-i360 ...	44	0.36
gi|13794006|gb|AY029007.1| Pinus taeda isolate PTLTP3-v94 n...	44	0.36
gi|13794003|gb|AY029005.1| Pinus taeda isolate PTLTP6-i425 ...	44	0.36
gi|24210408|emb|AJ320160.1|FNU320160 Fusobacterium nucleatu...	44	0.36
gi|13928026|emb|AL121594.6|CNS01DRY Human chromosome 14 DNA...	44	0.36
gi|8191116|gb|AC040163.3|AC040163 Homo sapiens chromosome 1...	44	0.36
gi|6707080|gb|AF139182.1|AF139182 Bartonella henselae filam...	44	0.36
gi|4028938|gb|AC004230.1|AC004230 Homo sapiens Chromosome 1...	44	0.36
gi|21623971|dbj|AP001094.6| Homo sapiens genomic DNA, chrom...	44	0.36
gi|19879812|dbj|AP001363.4| Homo sapiens genomic DNA, chrom...	44	0.36

Alignments

>gi|18868|emb|X13443.1|HVAATI Barley mRNA for alpha-amylase/trypsin inhibitor

Length = 672

Score = 81.8 bits (41), Expect = 2e-12
Identities = 41/41 (100%)
Strand = Plus / Plus

```
Query: 793 aagagattgaaccaacgaccaataaactagtatcaacaatg 833
           |||||||||||||||||||||||||||||||||||||||||
Sbjct: 1   aagagattgaaccaacgaccaataaactagtatcaacaatg 41
```

>gi|22035404|gb|AY124482.1| Danio rerio myogenin gene, exons 1, 2 and 3 and complete cds

Length = 4260

Score = 79.8 bits (40), Expect = 6e-12
Identities = 40/40 (100%)
Strand = Plus / Minus

```
Query: 3    tcgattactatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||||||||
Sbjct: 4246 tcgattactatagggcacgcgtggtcgacggcccgggctg 4207
```

>gi|24614584|gb|AY050653.1| Giardia intestinalis SEC24 (SEC24) mRNA, complete cds

Length = 1420

Score = 75.8 bits (38), Expect = 1e-10
Identities = 38/38 (100%)
Strand = Plus / Minus

```
Query: 5    gattactatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||||||
Sbjct: 1417 gattactatagggcacgcgtggtcgacggcccgggctg 1380
```

```
 Score = 67.9 bits (34), Expect = 2e-08
 Identities = 34/34 (100%)
 Strand = Plus / Plus


Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 23 actatagggcacgcgtggtcgacggcccgggctg 56
```

```
>gi|10938038|gb|AF162890.1|AF162890S1     Mus musculus peroxisomal assembly
protein PEX3P (Pex3) gene,
          promoter and exon 1
          Length = 2785

 Score = 75.8 bits (38), Expect = 1e-10
 Identities = 38/38 (100%)
 Strand = Plus / Plus


Query: 5  gattactatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||||||
Sbjct: 13 gattactatagggcacgcgtggtcgacggcccgggctg 50
```

```
>gi|5821237|dbj|AB021922.1|   Homo sapiens gene for lectin-like oxidized LDL
receptor, promoter
          region
          Length = 2463

 Score = 75.8 bits (38), Expect = 1e-10
 Identities = 38/38 (100%)
 Strand = Plus / Plus


Query: 5  gattactatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||||||
Sbjct: 13 gattactatagggcacgcgtggtcgacggcccgggctg 50
```

```
>gi|3721562|dbj|AB011276.1|     Mus musculus gene for alpha1a calcium
channel, partial cds
          Length = 2099

 Score = 73.8 bits (37), Expect = 4e-10
 Identities = 40/41 (97%)
 Strand = Plus / Plus


Query: 9  actatagggcacgcgtggtcgacggcccgggctgttattgg 49
          |||||||||||||||||||||||||||||||| ||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctggtattgg 41
```

>gi|30313388|gb|AY099112.1| Rattus norvegicus obese protein gene, 5' flanking region and
partial cds
Length = 1127

Score = 71.9 bits (36), Expect = 2e-09
Identities = 36/36 (100%)
Strand = Plus / Plus

```
Query: 7  ttactatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||||
Sbjct: 17 ttactatagggcacgcgtggtcgacggcccgggctg 52
```

>gi|18140057|gb|AF457660.1| Castanea dentata clone ACS2 vascular protein gene, promoter
region
Length = 685

Score = 71.9 bits (36), Expect = 2e-09
Identities = 36/36 (100%)
Strand = Plus / Plus

```
Query: 7  ttactatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||||
Sbjct: 9  ttactatagggcacgcgtggtcgacggcccgggctg 44
```

>gi|4104807|gb|AF039526.1|AF039526 Homo sapiens MHC class I related protein 1 (MR1) gene, partial
cds
Length = 1423

Score = 71.9 bits (36), Expect = 2e-09
Identities = 36/36 (100%)
Strand = Plus / Plus

```
Query: 7  ttactatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||||
Sbjct: 4  ttactatagggcacgcgtggtcgacggcccgggctg 39
```

>gi|8050595|gb|AF233737.1|AF233737 Agrotis ipsilon AiC5 chymotrypsinogen gene, promoter region
Length = 951

Score = 71.9 bits (36), Expect = 2e-09
Identities = 36/36 (100%)
Strand = Plus / Plus

```
Query: 7  ttactatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||||
Sbjct: 11 ttactatagggcacgcgtggtcgacggcccgggctg 46
```

>gi|6690643|gb|AF191544.1|AF191544 Homo sapiens estrogen receptor beta gene, promoter region and
partial cds
Length = 2495

Score = 71.9 bits (36), Expect = 2e-09
Identities = 39/40 (97%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctgttattg 48
          |||||||||||||||||||||||||||||||||||| |||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctggtattg 40
```

>gi|6164589|gb|AF051769.1|AF051769 Homo sapiens hyaluronidase-like pseudogene 1 (HYALP1), partial sequence
Length = 11957

Score = 71.9 bits (36), Expect = 2e-09
Identities = 36/36 (100%)
Strand = Plus / Minus

```
Query: 7      ttactatagggcacgcgtggtcgacggcccgggctg 42
               ||||||||||||||||||||||||||||||||||||
Sbjct: 11954 ttactatagggcacgcgtggtcgacggcccgggctg 11919
```

>gi|3643823|gb|AF075270.1|AF075270 Hordeum vulgare high affinity sulfate transporter (HVST1) gene,
promoter region, 5'UTR, and partial cds
Length = 1094

Score = 71.9 bits (36), Expect = 2e-09
Identities = 36/36 (100%)
Strand = Plus / Plus

```
Query: 7  ttactatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||||
Sbjct: 64 ttactatagggcacgcgtggtcgacggcccgggctg 99
```

>gi|19919401|gb|AF435445.1| Pleurotus ostreatus manganese peroxidase (mnp3) gene, promoter
region and partial cds
Length = 2790

Score = 69.9 bits (35), Expect = 6e-09
Identities = 35/35 (100%)
Strand = Plus / Plus

```
Query: 8   tactatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||
Sbjct: 79  tactatagggcacgcgtggtcgacggcccgggctg 113
```

>gi|8132114|gb|AF153014.1| Trichomonas vaginalis Tvp14 (tvp14) gene, complete cds
Length = 1463

Score = 69.9 bits (35), Expect = 6e-09
Identities = 35/35 (100%)
Strand = Plus / Plus

```
Query: 8  tactatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 11 tactatagggcacgcgtggtcgacggcccgggctg 45
```

>gi|4456992|gb|AF077743.1|AF077743 Mus musculus transcription factor TFEC gene, promoter region and
5' UTR
Length = 615

Score = 69.9 bits (35), Expect = 6e-09
Identities = 38/39 (97%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctgttatt 47
          ||||||||||||||||||||||||||||||||| ||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctggtatt 39
```

>gi|12276207|gb|AF269146.1|AF269146 Bilophila wadsworthia taurine:pyruvate aminotransferase gene,
complete cds
Length = 2050

Score = 69.9 bits (35), Expect = 6e-09
Identities = 35/35 (100%)
Strand = Plus / Minus

```
Query: 8    tactatagggcacgcgtggtcgacggcccgggctg 42
            ||||||||||||||||||||||||||||||||||
Sbjct: 2050 tactatagggcacgcgtggtcgacggcccgggctg 2016
```

>gi|14275833|emb|AJ289605.1|MMU289605 Mus musculus partial Lancl1 gene for LanC-like protein 1, exon 4
Length = 682

Score = 69.9 bits (35), Expect = 6e-09
Identities = 35/35 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctgt 43
          ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctgt 35
```

```
 Score = 67.9 bits (34), Expect = 2e-08
 Identities = 34/34 (100%)
 Strand = Plus / Plus


Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 46 actatagggcacgcgtggtcgacggcccgggctg 79



>gi|3378604|emb|AJ009889.1|HSAJ9889   Homo sapiens GHR gene, V1 promoter
region
          Length = 1640

 Score = 69.9 bits (35), Expect = 6e-09
 Identities = 38/39 (97%)
 Strand = Plus / Plus


Query: 9  actatagggcacgcgtggtcgacggcccgggctgttatt 47
          ||||||||||||||||||||||||||||||||| ||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctggtatt 39



>gi|3916231|gb|AF074397.1|AF074397   Homo sapiens anti-mullerian hormone type
II receptor (AMHR2)
          gene, promoter region and partial cds
          Length = 1135

 Score = 69.9 bits (35), Expect = 6e-09
 Identities = 38/39 (97%)
 Strand = Plus / Plus


Query: 9  actatagggcacgcgtggtcgacggcccgggctgttatt 47
          ||||||||||||||||||||||||||||||||| ||||
Sbjct: 13 actatagggcacgcgtggtcgacggcccgggctggtatt 51



>gi|5139506|emb|Z18892.2|MMDESMINP      Mus musculus desmin gene
          Length = 19391

 Score = 69.9 bits (35), Expect = 6e-09
 Identities = 35/35 (100%)
 Strand = Plus / Minus


Query: 8     tactatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||||
Sbjct: 19391 tactatagggcacgcgtggtcgacggcccgggctg 19357
```

>gi|15487305|dbj|AB060299.1| Mus musculus gene for acetyl CoA synthethase-1, promoter region, partial sequence
Length = 2094

Score = 69.9 bits (35), Expect = 6e-09
Identities = 38/39 (97%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctgttatt 47
          ||||||||||||||||||||||||||||||||| ||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctggtatt 39
```

>gi|12697590|dbj|AB046716.1| Homo sapiens hST3Gal I gene for alpha 2,3-sialyltransferase I, promoter region
Length = 1950

Score = 69.9 bits (35), Expect = 6e-09
Identities = 38/39 (97%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctgttatt 47
          ||||||||||||||||||||||||||||||||| ||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctggtatt 39
```

>gi|25453365|gb|AY050651.2| Giardia intestinalis MYB (MYB) mRNA, complete cds
Length = 3069

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Minus

```
Query: 9    actatagggcacgcgtggtcgacggcccgggctg 42
             ||||||||||||||||||||||||||||||||||
Sbjct: 3069 actatagggcacgcgtggtcgacggcccgggctg 3036
```

Score = 52.0 bits (26), Expect = 0.001
Identities = 26/26 (100%)
Strand = Plus / Plus

```
Query: 17 gcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||
Sbjct: 1  gcacgcgtggtcgacggcccgggctg 26
```

>gi|13183059|gb|AF237414.1| Anaplasma phagocytophilum RNA polymerase beta subunit (rpoB) gene, complete cds
Length = 4185

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Minus

```
Query: 9    actatagggcacgcgtggtcgacggcccgggctg 42
             ||||||||||||||||||||||||||||||||||
Sbjct: 4041 actatagggcacgcgtggtcgacggcccgggctg 4008
```

>gi|13625520|gb|AY014277.1| Lolium perenne gibberellin 20-oxidase gene, complete cds
Length = 2128

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||
Sbjct: 15 actatagggcacgcgtggtcgacggcccgggctg 48
```

>gi|17105179|gb|AF439558.1|AF439558 Mus musculus X2CR1 gene, promoter region and partial cds
Length = 830

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34
```

>gi|16209547|gb|AY052528.1| Glycine max calmodulin isoform-4 (cam-4) gene, promoter region
and partial cds
Length = 2050

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 5  gattactatagggcacgcgtggtcgacggcccgg 38
           ||||||||||||||||||||||||||||||||||
Sbjct: 13 gattactatagggcacgcgtggtcgacggcccgg 46
```

>gi|15213480|gb|AF241535.1|AF241535 Homo sapiens mucin 4 (MUC4) gene, promoter sequence and partial
cds
Length = 3716

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34
```

>gi|15558849|emb|AJ310936.1|FSO310936 Fusarium solani chy gene for putative cyanide hydratase enzyme
Length = 1981

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Minus

```
Query: 9    actatagggcacgcgtggtcgacggcccgggctg 42
            ||||||||||||||||||||||||||||||||||
Sbjct: 1981 actatagggcacgcgtggtcgacggcccgggctg 1948
```

>gi|4878023|gb|AF131239.2|AF131239 Rattus norvegicus alpha 1,2-fucosyltransferase C (FTC) gene, complete
cds
Length = 1555

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Minus

```
Query: 9    actatagggcacgcgtggtcgacggcccgggctg 42
            ||||||||||||||||||||||||||||||||||
Sbjct: 1534 actatagggcacgcgtggtcgacggcccgggctg 1501
```

>gi|15216031|emb|AJ318812.1|VFA318812 Vicia faba var. minor aap1 gene, promoter region
Length = 1702

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34
```

>gi|4454294|emb|AJ132779.1|MMU132779 Mus musculus p107 gene promoter region
Length = 776

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34
```

>gi|3133088|emb|Y16736.1|HSA16736 Homo sapiens dif-2 gene, promoter region
Length = 1368

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 11 actatagggcacgcgtggtcgacggcccgggctg 44
```

>gi|6690521|gb|AF154245.1|AF154245 Rattus norvegicus chemotactic protein-3 gene, complete cds
Length = 2416

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34
```

>gi|5731977|gb|AF114032.1|AF114032 Mus musculus glycogenin-1 gene, promoter and partial cds
Length = 1958

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34
```

>gi|26453412|dbj|AB094665.1| Seriola quinqueradiata YGHL1 gene for putative growth hormone
like protein-1, complete cds
Length = 6658

```
 Score = 67.9 bits (34), Expect = 2e-08
 Identities = 34/34 (100%)
 Strand = Plus / Plus


Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 39 actatagggcacgcgtggtcgacggcccgggctg 72


>gi|4039145|gb|AF099083.1|AF099083   Homo sapiens growth hormone secretagogue
receptor gene, 5'
          flanking region and partial cds
          Length = 1237

 Score = 67.9 bits (34), Expect = 2e-08
 Identities = 34/34 (100%)
 Strand = Plus / Plus


Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34


>gi|2739123|gb|AF029214.1|MMOX2S1      Mus musculus MRC OX-2 antigen homolog
gene, exon 1
          Length = 2791

 Score = 67.9 bits (34), Expect = 2e-08
 Identities = 34/34 (100%)
 Strand = Plus / Minus


Query: 9     actatagggcacgcgtggtcgacggcccgggctg 42
              ||||||||||||||||||||||||||||||||||
Sbjct: 2791 actatagggcacgcgtggtcgacggcccgggctg 2758


 Score = 67.9 bits (34), Expect = 2e-08
 Identities = 34/34 (100%)
 Strand = Plus / Plus


Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34


>gi|2895903|gb|AF046916.1|AF046916   Ruminococcus flavefaciens FD-1 glutamine
synthetase type III (glnA)
            gene, complete cds
          Length = 2685

 Score = 67.9 bits (34), Expect = 2e-08
 Identities = 34/34 (100%)
 Strand = Plus / Minus
```

```
Query: 9     actatagggcacgcgtggtcgacggcccgggctg 42
             ||||||||||||||||||||||||||||||||||
Sbjct: 2685 actatagggcacgcgtggtcgacggcccgggctg 2652
```

Score = 58.0 bits (29), Expect = 2e-05
Identities = 29/29 (100%)
Strand = Plus / Plus

```
Query: 14 agggcacgcgtggtcgacggcccgggctg 42
          ||||||||||||||||||||||||||||
Sbjct: 1  agggcacgcgtggtcgacggcccgggctg 29
```

>gi|1916583|gb|U53907.1|RNU53907 Rattus norvegicus microsatellite sequence D10Mco29
Length = 498

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||
Sbjct: 5  actatagggcacgcgtggtcgacggcccgggctg 38
```

>gi|6683556|dbj|AB024534.1| Rattus norvegicus gene for thiazide-sensitive Na-Cl
cotransporter, 5' flanking region
Length = 2145

Score = 67.9 bits (34), Expect = 2e-08
Identities = 34/34 (100%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||
Sbjct: 1  actatagggcacgcgtggtcgacggcccgggctg 34
```

>gi|7109286|gb|AF227508.1| Rattus norvegicus intestinal alkaline phosphatase-II (IAP-II) gene,
complete cds
Length = 6359

Score = 65.9 bits (33), Expect = 1e-07
Identities = 33/33 (100%)
Strand = Plus / Minus

```
Query: 7    ttactatagggcacgcgtggtcgacggcccggg 39
            |||||||||||||||||||||||||||||||||
Sbjct: 6349 ttactatagggcacgcgtggtcgacggcccggg 6317
```

>gi|13345792|gb|AF332759.1| Homo sapiens partially duplicated CHRNA7 gene, hybrid intron A/4
and partial exon 5
Length = 1280

Score = 65.9 bits (33), Expect = 1e-07
Identities = 36/37 (97%)
Strand = Plus / Plus

```
Query: 11 tatagggcacgcgtggtcgacggcccgggctgttatt 47
           |||||||||||||||||||||||||||||||| ||||
Sbjct: 8  tatagggcacgcgtggtcgacggcccgggctggtatt 44
```

>gi|5002511|emb|Z98266.1|HSZ98266 Homo sapiens gene encoding plakophilin (exons 1-13)
Length = 49999

Score = 65.9 bits (33), Expect = 1e-07
Identities = 33/33 (100%)
Strand = Plus / Plus

```
Query: 10 ctatagggcacgcgtggtcgacggcccgggctg 42
           |||||||||||||||||||||||||||||||||
Sbjct: 16 ctatagggcacgcgtggtcgacggcccgggctg 48
```

>gi|6560627|gb|AF112228.1|HSCD30P1 Homo sapiens CD30 protein (CD30) gene, promoter, partial sequence
Length = 1605

Score = 65.9 bits (33), Expect = 1e-07
Identities = 37/39 (94%)
Strand = Plus / Plus

```
Query: 9  actatagggcacgcgtggtcgacggcccgggctgttatt 47
           ||||||||||||||||||||||||||| ||||| ||||
Sbjct: 1  actatagggcacgcgtggtcgacggccckggctggtatt 39
```

>gi|24475537|dbj|AB084484.1| Betula platyphylla DNA, microsatellite:BpA
Length = 427

Score = 65.9 bits (33), Expect = 1e-07
Identities = 33/33 (100%)
Strand = Plus / Minus

```
Query: 10  ctatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||||
Sbjct: 317 ctatagggcacgcgtggtcgacggcccgggctg 285
```

>gi|28875405|gb|AF515463.1| Biomphalaria glabrata fibrinogen related protein 12.1 precursor (FREP12.1) gene, partial cds
Length = 609

Score = 63.9 bits (32), Expect = 4e-07
Identities = 32/32 (100%)
Strand = Plus / Plus

```
Query: 11 tatagggcacgcgtggtcgacggcccgggctg 42
           ||||||||||||||||||||||||||||||||
Sbjct: 1  tatagggcacgcgtggtcgacggcccgggctg 32
```

>gi|18140058|gb|AF457661.1| Castanea dentata clone ACS10A vascular protein gene, promoter region
Length = 1157

Score = 63.9 bits (32), Expect = 4e-07
Identities = 35/36 (97%)
Strand = Plus / Plus

```
Query: 7  ttactatagggcacgcgtggtcgacggcccgggctg 42
           |||||||||||||||||||||| ||||||||||||
Sbjct: 10 ttactatagggcacgcgtggtcgccggcccgggctg 45
```

>gi|2581766|gb|U77633.1|RNU77633 Rattus norvegicus chromosome 10 microsatellite D10Mco34
Length = 1023

Score = 63.9 bits (32), Expect = 4e-07
Identities = 33/34 (97%)
Strand = Plus / Minus

```
Query: 9    actatagggcacgcgtggtcgacggcccgggctg 42
             |||||||||||||| ||||||||||||||||||
Sbjct: 1023 actatagggcacgcvtggtcgacggcccgggctg 990
```

Database: All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS, GSS, or phase 0, 1 or 2 HTGS sequences)

PCT/GB2004/002589